

Zależność między miejscem zamieszkania na wsi i w mieście a głównym źródłem utrzymania ludności

Praca najemna jest wciąż najczęstszym źródłem utrzymania¹ ludności w Polsce. Jednak liczba ludności, dla której dochody z pracy najemnej są głównym źródłem utrzymania nie stanowi nawet połowy liczebności mieszkańców kraju mających własne źródło utrzymania, zatem nie jest jedynym ważnym źródłem utrzymania Polaków. W 2011 r. głównie z pracy najemnej utrzymywało się 48,9% osób mających własne źródło dochodów, z pracy na własny rachunek poza rolnictwem — 6,1%, z pracy na własny rachunek w rolnictwie — 4,8%, ze źródeł niezarobkowych — 40,0%, a z własności — niecałe 0,2%. Z kolei

¹ Przy przeprowadzaniu badań Autorki posłużyły się klasyfikacją źródeł utrzymania podaną przez GUS. Zgodnie z tą klasyfikacją:

- **praca najemna** obejmuje zarówno pracę świadczoną w sektorze publicznym, jak i pracę świadczoną w sektorze prywatnym;
- **praca na rachunek własny poza rolnictwem** to prowadzenie własnej działalności gospodarczej albo wykonywanie wolnego zawodu. W opisywanej grupie znajdują się również osoby prowadzące jednoosobowy podmiot gospodarczy, którym to osobom narzucono zarejestrowanie własnej działalności gospodarczej. Co istotne, także dochody z wynajmu są traktowane w wykorzystywanej tutaj klasyfikacji jako pochodzące z prowadzonej działalności gospodarczej, stąd osoby, dla których głównym źródłem dochodu jest wynajem, w prezentowanych zestawieniach nie zostały ujęte oddzielnie, lecz razem z pracującymi na własny rachunek poza rolnictwem;
- **praca na rachunek własny w rolnictwie** to praca w swoim gospodarstwie rolnym związana z produkcją roślinną i/lub zwierzęcą. Do grupy tej zaliczono również członków rolniczych spółdzielni produkcyjnych, kółek rolniczych oraz członków spółek cywilnych w rolnictwie. Jako praca na własny rachunek w rolnictwie traktowana jest również praca wykonywana przez osoby utrzymujące się z prowadzenia działalności usługowej związanej z rolnictwem (czyli taka działalność, jak: nawożenie pól, opryski upraw, obsługa systemów irygacyjnych, sztuczne unasienianie zwierząt, usługi pasterskie, usługi w zakresie leśnictwa lub łowiectwa);
- **niezarobkowe źródło dochodu** to emerytura (pracownicza, kombatancka, rolna), renta: strukturalna, z tytułu niezdolności do pracy, inwalidzka, rodzinna (wdowia, sieroca) i socjalna, zasiłek dla bezrobotnych, świadczenia i zasiłki przedemerytalne oraz świadczenia pomocy społecznej;
- **dochody z własności** to dochody z tytułu oddanych w dzierżawę gruntów rolnych, dochody z lokat kapitałowych (z obrotu akcjami, obligacjami, z zysków kapitałowych itp.), z odsetek od oszczędności i z udziału w zyskach przedsiębiorstw (dywidend) — *Ludność...* (2013), s. 22—24.

Dodatkowo GUS w klasyfikacji podaje też **inne źródła niż wymienione** (do kategorii tej należą m.in. alimenty od osób prywatnych, stypendia naukowe i sportowe), jednak znaczenie tej pozycji — w przypadku rozpatrywania w niniejszym artykule jedynie osób posiadających własne źródło utrzymania — jest znikome i dlatego też we wszystkich przeprowadzonych analizach zostało pominięte.

w 2002 r. głównie z pracy najemnej utrzymywało się w Polsce 41,7% osób mających własne źródło dochodów, z pracy na własny rachunek poza rolnictwem — 5,5%, z pracy na własny rachunek w rolnictwie — 6,4%, ze źródeł niezarobkowych — 46,3%, a z własności — jedynie nieco ponad 0,1%. Z porównania przywołanych danych za 2011 r. z odpowiednimi danymi dotyczącymi 2002 r. można wyciągnąć następujące wnioski²:

- znacznie wzrósł udział osób, dla których dochody z pracy najemnej są głównym źródłem utrzymania (wzrost o 7,2 p.proc.);
- wzrósł udział osób, w przypadku których dochody z pracy na własny rachunek poza rolnictwem są głównym źródłem utrzymania (wzrost o 0,6 p.proc.), a spadł udział osób, dla których dochody z pracy na własny rachunek w rolnictwie są głównym źródłem utrzymania (spadek o 1,6 p.proc.). Dzięki tym zmianom w 2011 r. więcej mieszkańców Polski czerpało główne dochody z pracy na własny rachunek poza rolnictwem niż z pracy na własny rachunek w rolnictwie, podczas gdy w 2002 r. było na odwrót;
- udział osób, w przypadku których dochody ze źródeł niezarobkowych są głównym źródłem utrzymania spadł o 6,3 p.proc.;
- z własności swoje główne dochody pobierało w 2011 r. więcej osób niż w 2002 r. i liczebność tej grupy rosła najbardziej dynamicznie (liczba takich osób zwiększyła się na przestrzeni dziewięciu lat o 32%). Jednak udział osób, dla których dochody z własności były głównym źródłem utrzymania w 2002 r. był bardzo niski, z tego względu — pomimo dynamicznego wzrostu na przestrzeni badanego okresu — udział ten w 2011 r. nadal pozostał relatywnie niski i zwiększył się jedynie o część p.proc.

Na to, jak kształtuje się struktura ludności pod względem głównego źródła utrzymania wpływ ma wiele różnorodnych czynników³. Czynnikiem, który kształtuje strukturę ludności Polski — w przypadku rozpatrywania głównego źródła utrzymania Polaków jako cechy będącej przedmiotem badania — jest także miejsce zamieszkania. I właśnie odpowiedź na pytanie, w jakim stopniu fakt, że osoba mieszka na wsi lub w mieście wpływa na rodzaj jej głównego źró-

² Wszystkie wnioski oparto na danych pochodzących z Narodowego Spisu Powszechnego Ludności i Mieszkań 2002 oraz Narodowego Spisu Powszechnego Ludności i Mieszkań 2011, których wyniki Główny Urząd Statystyczny udostępnił w publikacjach *Ludność...* (2003), s. 62—79 oraz *Ludność...* (2013), s. 92—97.

³ Takimi czynnikami są przykładowo wiek i wykształcenie. Zdecydowanie większy jest udział osób utrzymujących się z pracy najemnej w przypadku trzydziesto- czy czterdziestolatków niż wśród siedemdziesięcio- i osiemdziesięciolatków. Z kolei udział osób utrzymujących się ze źródeł niezarobkowych jest większy wśród siedemdziesięcio- i osiemdziesięciolatków niż w grupie trzydziesto- czy czterdziestolatków — *Ludność...* (2013), s. 32—34. Czynnikiem, który ma wpływ na strukturę społeczeństwa pod względem głównego źródła utrzymania jest również wykształcenie. Okazuje się, że osoby ze średnim i wyższym wykształceniem relatywnie częściej czerpią swoje główne dochody z pracy najemnej bądź z pracy na własny rachunek poza rolnictwem niż osoby z wykształceniem zasadniczym zawodowym, gimnazjalnym czy podstawowym. Z kolei dla osób z wykształceniem podstawowym, gimnazjalnym i zasadniczym zawodowym stosunkowo częściej źródła niezarobkowe są głównym źródłem utrzymania — *Ludność...* (2013), s. 36—39.

dła dochodu, stała się celem obecnego opracowania. W związku z celem pracy postawiono hipotezę, że fakt mieszkania w mieście albo na wsi w coraz mniejszym stopniu determinuje rozkład mieszkańców Polski pod względem rodzajów posiadanych źródeł utrzymania. Aby udzielić pełnych odpowiedzi dotyczących tych kwestii, przeprowadzono trzyetapowe badanie, które stanowi kolejne podrozdziały tego opracowania. Badaniem objęto następujące zagadnienia:

- 1) określenie związku między miejscem zamieszkania a głównym źródłem utrzymania mieszkańców Polski w 2002 r.;
- 2) określenie związku między miejscem zamieszkania a głównym źródłem utrzymania mieszkańców Polski w 2011 r.;
- 3) porównanie siły zależności między miejscem zamieszkania a głównym źródłem utrzymania mieszkańców Polski w latach 2002 i 2011.

W celu sprawdzenia, czy występuje statystycznie istotny związek między rozpatrywanymi cechami wykorzystano test niezależności *chi*-kwadrat przeprowadzony dla lat 2002 i 2011. Z kolei do określenia siły zależności występujących w poszczególnych latach użyto współczynnika *V* Cramera. W ostatnim etapie — na podstawie porównania wartości współczynnika *V* Cramera obliczonego dla 2002 r. z jego wartością uzyskaną dla 2011 r. — wyciągnięto wnioski pozwalające na całkowitą realizację głównego celu badania.

ISTOTA ZASTOSOWANEGO NARZĘDZIA BADAWCZEGO

Test niezależności *chi*-kwadrat jest nieparametrycznym testem istotności, który może być wykorzystywany do oceny zależności stochastycznej dwóch cech jakościowych, dwóch cech ilościowych, a także dowolnej cechy ilościowej i jakościowej. Punktem wyjścia do przeprowadzenia testu jest sporządzenie tablicy, której wnętrze stanowi liczebność empiryczna, czyli zaobserwowana. Tablica ta jest macierzą o *r* wierszach i *s* kolumnach, przy czym *r* oznacza liczbę wariantów pierwszej cechy (*X*), a *s* — liczbę wariantów drugiej cechy (*Y*). Z kolei n_{ij} ($i=1, 2, \dots, r, j=1, 2, \dots, s$) to liczba tych obserwacji, dla których cecha *X* przyjmuje wariant x_i , a cecha *Y* — wariant y_j .

Sumując osobno każdy wiersz i każdą kolumnę macierzy liczebności empirycznej otrzymuje się tzw. liczebność brzegową, którą oznaczono jako n_i i n_j . Zachodzą więc równości⁴:

$$n_i = \sum_{j=1}^s n_{ij} \quad n_j = \sum_{i=1}^r n_{ij} \quad (1)$$

Czyli n_i to liczebność brzegowa w wierszu o numerze *i*, którą uzyskano dodając każdą liczebność znajdującą się w tym wierszu, natomiast n_j to liczebność brzegowa w kolumnie o numerze *j*, którą obliczono dodając każdą liczeb-

⁴ Kot i in. (2007), s. 294 i 295.

ność leżącą w tej kolumnie. Poprawność wykonanych obliczeń można skontrolować sprawdzając, czy suma liczebności brzegowej dotycząca wierszy jest taka sama, jak suma liczebności brzegowej z kolumn i równa się liczebności całkowitej n . Tak więc:

$$n = \sum_{i=1}^r \sum_{j=1}^s n_{ij} = \sum_{i=1}^r n_{i\cdot} = \sum_{j=1}^s n_{\cdot j} \quad (2)$$

Tabl. 1 przedstawia schemat macierzy liczebności empirycznej.

TABL. 1. SCHEMAT MACIERZY LICZEBNOŚCI EMPIRYCZNEJ

Wyszczególnienie		Warianty drugiej cechy (Y)						$n_{i\cdot}$
		y_1	y_2	...	y_j	...	y_s	
Warianty pierwszej cechy (X)	x_1	n_{11}	n_{12}	...	n_{1j}	...	n_{1s}	$n_{1\cdot}$
	x_2	n_{21}	n_{22}	...	n_{2j}	...	n_{2s}	$n_{2\cdot}$

	x_i	n_{i1}	n_{i2}	...	n_{ij}	...	n_{is}	$n_{i\cdot}$

	x_r	n_{r1}	n_{r2}	...	n_{rj}	...	n_{rs}	$n_{r\cdot}$
$n_{\cdot j}$		$n_{\cdot 1}$	$n_{\cdot 2}$...	$n_{\cdot j}$...	$n_{\cdot s}$	n

Źródło: opracowanie własne, gdzie: r — liczba wariantów cechy X , s — liczba wariantów cechy Y , n_{ij} — liczba obserwacji posiadających i -ty wariant cechy X oraz j -ty wariant cechy Y .

Mając tak przygotowaną macierz liczebności empirycznej można sformułować hipotezę zerową H_0 , iż badane dwie cechy są stochastycznie niezależne wobec hipotezy alternatywnej H_1 , ponieważ występuje stochastyczna zależność między tymi cechami⁵. Aby sprawdzić prawdziwość hipotezy zerowej, na podstawie obliczonej liczebności brzegowej i liczebności całkowitej należy wyznaczyć prawdopodobieństwa brzegowe zgodnie z formułami⁶:

$$p_{i\cdot} = \frac{n_{i\cdot}}{n} \quad p_{\cdot j} = \frac{n_{\cdot j}}{n} \quad (3)$$

gdzie:

- $p_{i\cdot}$ — prawdopodobieństwo, że obserwacja przyjmuje i -ty wariant cechy X ,
- $p_{\cdot j}$ — prawdopodobieństwo, że obserwacja przyjmuje j -ty wariant cechy Y .

⁵ Aczel (2000), s. 758.

⁶ Kukuła (2003), s. 196.

Z kolei prawdopodobieństwa empiryczne wewnątrz tablicy można obliczyć według wzoru:

$$p_{ij} = \frac{n_{ij}}{n} \quad (4)$$

gdzie p_{ij} — prawdopodobieństwo, że obserwacja przyjmuje i -ty wariant cechy X i j -ty wariant cechy Y .

Następnie, mnożąc kolejne prawdopodobieństwa brzegowe dotyczące wierszy przez prawdopodobieństwa brzegowe z poszczególnych kolumn otrzymuje się macierz prawdopodobieństw teoretycznych p_{ij}^* . Obliczone wielkości są hipotetycznymi prawdopodobieństwami, które wystąpiłyby, gdyby hipoteza zerowa była prawdziwa, czyli gdyby rozpatrywane cechy X i Y były niezależne. Zatem prawdopodobieństwa teoretyczne wyznacza się zgodnie ze wzorem⁷:

$$p_{ij}^* = p_{i \cdot} \cdot p_{\cdot j} \text{ przy czym } \sum_{i=1}^r \sum_{j=1}^s p_{ij}^* = 1, \text{ czyli } 100\% \quad (5)$$

Treść hipotezy zerowej i hipotezy alternatywnej można formalnie przedstawić następująco⁸:

$$H_0: E(p_{ij} - p_{ij}^*)^2 = 0 \text{ przeciwko } H_1: E(p_{ij} - p_{ij}^*)^2 > 0 \quad (6)$$

gdzie symbol E oznacza wartość oczekiwaną.

W następnym etapie należy obliczyć liczebność hipotetyczną n_{ij}^* , która wystąpiłaby, gdyby spełniony był warunek o niezależności cech. Poszczególne wartości n_{ij}^* otrzymano mnożąc odpowiednie prawdopodobieństwa teoretyczne p_{ij}^* przez liczebność całkowitą n , a więc postępując według wzoru⁹:

$$n_{ij}^* = np_{ij}^* \quad (7)$$

⁷ Ostasiewicz i in. (1995), s. 264.

⁸ Podgórski (2005), s. 274.

⁹ Zeliaś (2000), s. 286.

Ostatecznie na podstawie porównania elementów macierzy liczebności rzeczywistej n_{ij} z elementami macierzy liczebności teoretycznej n_{ij}^* należy zdecydować, czy odrzucić hipotezę H_0 na skutek wystąpienia dużych różnic między tymi dwoma rodzajami liczebności czy jednak nie ma podstaw do odrzucenia hipotezy zerowej. Podjęcie decyzji ułatwia zastosowanie statystyki χ^2 , którą przyjmuje się jako syntetyczną miarę odchyłeń liczebności rzeczywistej od liczebności teoretycznej. Wartość statystyki testowej χ^2 obliczono zgodnie z formułą¹⁰:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - n_{ij}^*)^2}{n_{ij}} \quad (8)$$

Statystyka χ^2 , przy założeniu prawdziwości hipotezy H_0 o niezależności cech, ma asymptotyczny rozkład *chi*-kwadrat o $(r-1)(s-1)$ stopniach swobody. Przyjmuje ona wartości z przedziału $\langle 0, n \cdot \min(r-1)(s-1) \rangle$. Z jej budowy wynika, że im większe są rozbieżności między liczebnością empiryczną n_{ij} i oczekiwaną n_{ij}^* , tym wyższa jest wartość obliczonej statystyki χ^2 . Statystyka χ^2 jest równa zero, gdy poszczególna liczebność empiryczna i teoretyczna są takie same, a zatem rozpatrywane cechy są stochastycznie niezależne. Z kolei maksymalną wartość, tj. $n \cdot \min(r-1)(s-1)$, przyjmuje ona w przypadku zależności funkcyjnej między cechami X i Y ¹¹.

Przy podejmowaniu decyzji w teście niezależności *chi*-kwadrat należy brać pod uwagę jednostronny (a konkretnie — prawostronny) obszar krytyczny, który określa nierówność $\chi^2 \geq \chi_{\alpha}^2$, gdzie χ_{α}^2 jest wartością krytyczną odczytaną z tablic rozkładu *chi*-kwadrat dla przyjętego z góry poziomu istotności α i dla $(r-1)(s-1)$ stopni swobody w taki sposób, aby zachodziła relacja $P\{\chi^2 \geq \chi_{\alpha}^2\} = \alpha$. Obliczoną wartość statystyki testowej χ^2 porównuje się z wartością krytyczną χ_{α}^2 i jeśli spełniona jest nierówność $\chi^2 \geq \chi_{\alpha}^2$, to przy przyjętym poziomie istotności hipotezę H_0 należy odrzucić na korzyść hipotezy H_1 . Oznacza to, że rozpatrywane cechy są zależne. Gdy natomiast zachodzi nierówność $\chi^2 < \chi_{\alpha}^2$, nie ma podstaw do odrzucenia hipotezy zerowej o niezależności badanych cech.

Do określenia siły zależności między cechami wykorzystano współczynnik V Cramera. Współczynnik ten jest wielkością niemianowaną i unormowaną

¹⁰ Kryszicki i in. (2003), s. 104.

¹¹ Józwiak, Podgórski (1995), s. 233.

— przyjmuje wartości wyłącznie z przedziału $(0,1)$ ¹². Jeżeli wynosi 0, to między cechami nie występuje zależność. Z kolei im jego wartość jest bliższa 1, tym zależność jest silniejsza. Współczynnik V Cramera obliczono według wzoru¹³:

$$V = \sqrt{\frac{\chi^2}{n \cdot \min(r-1)(s-1)}} \quad (9)$$

ANALIZA WPŁYWU MIEJSCA ZAMIESZKANIA NA ROZKŁAD GŁÓWNYCH ŹRÓDEŁ UTRZYMANIA LUDNOŚCI W POLSCE W 2002 R.

Pierwszym postawionym przed Autorkami zadaniem jest odpowiedź na pytanie, czy w 2002 r. istniała zależność między tym, że osoba mieszka w mieście czy na wsi a rodzajem jej głównego źródła utrzymania. A zatem weryfikacji podlega hipoteza H_0 stanowiąca, że badane dwie cechy są stochastycznie niezależne wobec hipotezy alternatywnej H_1 orzekającej, że występuje stochastyczna zależność między tymi cechami. Procedurę weryfikacyjną zrealizowano za pomocą testu niezależności *chi*-kwadrat. Dane stanowiące punkt wyjścia do przeprowadzenia testu zaprezentowano w tabl. 2¹⁴.

TABL. 2. LICZEBNOŚĆ EMPIRYCZNA (rzeczywista) W 2002 R.

Wyszczególnienie	Główne dochody					Suma
	z pracy			ze źródeł niezarobkowych	z własności	
	najemnej	na własny rachunek poza rolnictwem	na własny rachunek w rolnictwie			
Miasta	6778680	952481	82260	6596832	19841	14430094
Wieś	2836644	311780	1392704	4094853	7229	8643210
Suma	9615324	1264261	1474964	10691685	27070	23073304

Źródło: opracowanie własne na podstawie publikacji *Ludność...* (2003), s. 68 i 74.

Na podstawie informacji prezentowanych w tabl. 2 oraz wykorzystując wzór (3) i wzór (4) wyznaczono prawdopodobieństwa empiryczne (tabl. 3).

¹² Pułaska-Turyńska (2005), s. 254.

¹³ Buga, Kassyk-Rokicka (2008), s. 121.

¹⁴ Jak już wspomniano, przeprowadzone badanie obejmuje jedynie tych mieszkańców Polski, którzy mają własne źródło utrzymania. Do badanej zbiorowości nie zakwalifikowały się zatem te osoby, które są na utrzymaniu innych osób. Zgodnie z danymi GUS w 2002 r. główne źródło utrzymania w postaci pracy najemnej, pracy na własny rachunek w rolnictwie bądź poza rolnictwem, własności albo źródła niezarobkowego miały 23073304 osoby.

TABL. 3. PRAWDOPODOBIEŃSTWA EMPIRYCZNE (rzeczywiste) W 2002 R. W %

Wyszczególnienie	Główne dochody					Suma
	z pracy			ze źródeł niezarobko- wych	z własności	
	najemnej	na własny rachunek poza rolnictwem	na własny rachunek w rolnictwie			
Miasta	29,379	4,128	0,357	28,591	0,086	62,540
Wieś	12,294	1,351	6,036	17,747	0,031	37,460
Suma	41,673	5,479	6,393	46,338	0,117	100,000

Źródło: opracowanie własne na podstawie tabl. 2.

Następnie obliczono prawdopodobieństwa teoretyczne, które występowałyby przy stochastycznej niezależności badanych cech. Do wyznaczenia tych prawdopodobieństw posłużył wzór (5). Otrzymane wyniki prezentuje tabl. 4.

TABL. 4. PRAWDOPODOBIEŃSTWA HIPOTETYCZNE (teoretyczne) W 2002 R. W %

Wyszczególnienie	Główne dochody					Suma
	z pracy			ze źródeł niezarobko- wych	z własności	
	najemnej	na własny rachunek poza rolnictwem	na własny rachunek w rolnictwie			
Miasta	26,062	3,427	3,998	28,980	0,073	62,540
Wieś	15,611	2,053	2,395	17,358	0,044	37,460
Suma	41,673	5,480	6,393	46,338	0,117	100,000

Źródło: opracowanie własne na podstawie tabl. 3.

Z kolei mnożąc — zgodnie ze wzorem (7) — wyznaczone prawdopodobieństwa hipotetyczne przez liczbę mieszkańców Polski, którzy w 2002 r. mieli własne źródło utrzymania (23073304), otrzymano liczebność hipotetyczną (tabl. 5).

TABL. 5. LICZEBNOŚĆ HIPOTETYCZNA (teoretyczna) W 2002 R.

Wyszczególnienie	Główne dochody					Suma
	z pracy			ze źródeł niezarobko- wych	z własności	
	najemnej	na własny rachunek poza rolnictwem	na własny rachunek w rolnictwie			
Miasta	6013444	790672	922446	6686603	16930	14430094
Wieś	3601880	473589	552518	4005082	10140	8643210
Suma	9615324	1264261	1474964	10691685	27070	23073304

Źródło: opracowanie własne na podstawie tabl. 2 i 4.

Dysponując wszystkimi elementami macierzy liczebności empirycznej n_{ij} (tabl. 2) oraz macierzy liczebności teoretycznej n_{ij}^* (tabl. 5) można wyznaczyć poszczególne składniki statystyki testowej χ^2 . Obliczona — na podstawie wzoru (8) — wartość statystyki χ^2 wyniosła 2395797. Czcionką pogrubioną zaznaczono największe wartości jej składników (tabl. 6).

TABL. 6. SKŁADNIKI STATYSTYKI TESTOWEJ CHI-KWADRAT W 2002 R.

Wyszczególnienie	Główne dochody					Suma
	z pracy			ze źródeł niezarobkowych	z własności	
	najmniej	na własny rachunek poza rolnictwem	na własny rachunek w rolnictwie			
Miasta	97379	33114	765261	1205	501	897461
Wieś	162578	55285	1277626	2012	836	1498337
Suma	259957	88399	2042887	3217	1337	2395798

Źródło: opracowanie własne na podstawie tabl. 2 i 5.

Mając obliczoną wartość χ^2 można przystąpić do weryfikacji hipotezy zerowej o niezależności cech będących przedmiotem analizy. Ponieważ $r=2$ oraz $s=5$, to liczba stopni swobody jest równa $(r-1)(s-1)=1 \cdot 4=4$. Jeśli przyjmie się poziom istotności $\alpha=0,001$, to dla 4 stopni swobody odczytana z tablic rozkładu chi-kwadrat wartość krytyczna χ^2_{α} wynosiła 18,5. Porównując obliczoną wartość χ^2 z odpowiednią wartością krytyczną otrzymano $\chi^2 = 2395797 > 18,5 = \chi^2_{\alpha}$.

Skoro uzyskano nierówność $\chi^2 > \chi^2_{\alpha}$, to hipotezę H_0 o niezależności głównego źródła utrzymania mieszkańca Polski w 2002 r. od jego miejsca zamieszkania należy odrzucić na rzecz hipotezy alternatywnej, orzekając tym samym, że rozpatrywane cechy są zależne. Można zatem stwierdzić, że istnieje statystycznie istotny związek między częstotliwością występowania wymienionych źródeł dochodu a tym, czy osoba mieszka w mieście czy na wsi. Okazało się bowiem, że uzyskane odchylenia między liczebnością empiryczną i teoretyczną były wystarczająco duże, aby odrzucić przypuszczenie o niezależności. Należy się jednak liczyć z możliwością popełnienia błędu odrzucenia hipotezy H_0 , pomimo tego że w rzeczywistości jest ona prawdziwa — prawdopodobieństwo popełnienia takiego błędu wynosi w tym przypadku 0,001. W teorii statystyki błąd ten nosi nazwę błędu I rodzaju.

Po przeprowadzeniu porównania poszczególnych prawdopodobieństw empirycznych znajdujących się w tabl. 3 z odpowiadającymi im prawdopodobieństwami teoretycznymi znajdującymi się w tabl. 4 można wyciągnąć wniosek, że

w 2002 r. osoby mieszkające w mieście względnie częściej czerpały dochody z pracy najemnej oraz z pracy na własny rachunek poza rolnictwem, natomiast osoby mieszkające na wsi relatywnie rzadziej. Z kolei osoby mieszkające na wsi relatywnie częściej osiągały dochody z pracy na własny rachunek w rolnictwie.

*ANALIZA WPLYWU MIEJSCA ZAMIESZKANIA NA ROZKŁAD
GŁÓWNYCH ŹRÓDEŁ UTRZYMANIA LUDNOŚCI W POLSCE W 2011 R.*

Drugim postawionym przed Autorkami zadaniem jest odpowiedź na pytanie, czy w 2011 r. istniała zależność między tym, czy osoba mieszka w mieście czy na wsi a tym, jakie jest jej główne źródło utrzymania. Ponownie weryfikacji będzie podlegała hipoteza H_0 stanowiąca, że badane dwie cechy są stochastycznie niezależne, wobec hipotezy alternatywnej H_1 orzekającej, że występuje stochastyczna zależność między tymi cechami. Dane niezbędne do przeprowadzenia weryfikacji za pomocą testu niezależności chi-kwadrat zebrano w tabl. 7¹⁵.

TABL. 7. LICZEBNOŚĆ EMPIRYCZNA (rzeczywista) W 2011 R.

Wyszczególnienie	Główne dochody					Suma
	z pracy			ze źródeł niezarobkowych	z własności	
	najemnej	na własny rachunek poza rolnictwem	na własny rachunek w rolnictwie			
Miasta	7914764	1040088	97828	6117058	19500	15189238
Wieś	3986774	455084	1059658	3625777	16242	9143535
Suma	11901538	1495172	1157486	9742835	35742	24332773

Źródło: opracowanie własne na podstawie publikacji *Ludność...* (2013).

Na podstawie informacji zaprezentowanych w tabl. 7 obliczono poszczególne prawdopodobieństwa empiryczne. Następnie, zakładając stochastyczną niezależność cech, obliczono prawdopodobieństwa teoretyczne. Z kolei mnożąc wyznaczone prawdopodobieństwa hipotetyczne przez liczbę mieszkańców Polski w 2011 r. posiadających własne źródło utrzymania (24332773), otrzymano liczebność hipotetyczną. Mając wszystkie elementy macierzy liczebności empirycznej n_{ij} i macierzy liczebności teoretycznej n_{ij}^* , wyznaczono wartość statystyki testowej χ^2 . Statystyka ta wyniosła w tym przypadku 1555791.

W tabl. 8—11 przedstawiono wyniki uzyskane ze wszystkich opisanych kroków obliczeniowych. W tabl. 11 czcionką pogrubioną zaznaczono największe składniki statystyki testowej χ^2 .

¹⁵ Zgodnie z danymi GUS w 2011 r. 24332773 mieszkańców Polski miało własne źródło utrzymania, a ich głównym źródłem utrzymania była: praca najemna, praca na własny rachunek w rolnictwie, praca na własny rachunek poza rolnictwem, własność albo źródło niezarobkowe.

TABL. 8. PRAWDOPODOBIENSTWA EMPIRYCZNE (rzeczywiste) W 2011 R. W %

Wyszczególnienie	Główne dochody					Suma
	z pracy			ze źródeł niezarobkowych	z własności	
	najmniej	na własny rachunek poza rolnictwem	na własny rachunek w rolnictwie			
Miasta	32,527	4,274	0,402	25,139	0,080	62,423
Wieś	16,384	1,870	4,355	14,901	0,067	37,577
Suma	48,911	6,144	4,757	40,040	0,147	100,000

Źródło: opracowanie własne na podstawie tabl. 7.

TABL. 9. PRAWDOPODOBIENSTWA HIPOTETYCZNE (teoretyczne) W 2011 R. W %

Wyszczególnienie	Główne dochody					Suma
	z pracy			ze źródeł niezarobkowych	z własności	
	najmniej	na własny rachunek poza rolnictwem	na własny rachunek w rolnictwie			
Miasta	30,532	3,836	2,969	24,994	0,092	62,423
Wieś	18,380	2,309	1,788	15,046	0,055	37,577
Suma	48,912	6,145	4,757	40,040	0,147	100,000

Źródło: opracowanie własne na podstawie tabl. 8.

TABL. 10. LICZEBNOŚĆ HIPOTETYCZNA (teoretyczna) W 2011 R.

Wyszczególnienie	Główne dochody					Suma
	z pracy			ze źródeł niezarobkowych	z własności	
	najmniej	na własny rachunek poza rolnictwem	na własny rachunek w rolnictwie			
Miasta	7429293	933331	722537	6081766	22311	15189238
Wieś	4472245	561841	434949	3661069	13431	9143535
Suma	11901538	1495172	1157486	9742835	35742	24332773

Źródło: opracowanie własne na podstawie tabl. 7 i 9.

TABL. 11. SKŁADNIKI STATYSTYKI TESTOWEJ CHI-KWADRAT W 2011 R.

Wyszczególnienie	Główne dochody					Suma
	z pracy			ze źródeł niezarobkowych	z własności	
	najmniej	na własny rachunek poza rolnictwem	na własny rachunek w rolnictwie			
Miasta	31723	12211	540126	205	354	584620
Wieś	52699	20285	897258	340	588	971171
Suma	84422	32496	1437384	545	942	1555791

Źródło: opracowanie własne na podstawie tabl. 7 i 10.

Skoro wyznaczono wartość statystyki χ^2 , to można przystąpić do weryfikacji hipotezy zerowej o niezależności cech będących przedmiotem analizy. Wartość krytyczna χ^2_α odczytana z tablic rozkładu chi-kwadrat dla poziomu istotności α równego 0,001 i 4 stopni swobody wynosi 18,5. Porównując obliczoną wartość statystyki χ^2 z wartością krytyczną uzyskano $\chi^2 = 1555791 > 18,5 = \chi^2_\alpha$.

Spełniona została nierówność $\chi^2 > \chi^2_\alpha$, wobec czego hipotezę H_0 o niezależności rozkładu głównych źródeł utrzymania mieszkańców Polski w 2011 r. od miejsca zamieszkania należy odrzucić na rzecz hipotezy alternatywnej. Można więc stwierdzić, że istnieje statystycznie istotny związek między częstotliwością występowania wymienionych pięciu źródeł utrzymania a tym, czy dana osoba mieszka na wsi czy w mieście. Należy się jednak liczyć z możliwością popełnienia błędu odrzucenia hipotezy H_0 , pomimo tego że w rzeczywistości jest ona prawdziwa — prawdopodobieństwo popełnienia takiego błędu wynosi 0,001. Jest to błąd I rodzaju.

Z porównania poszczególnych prawdopodobieństw empirycznych pokazanych w tabl. 8 z odpowiadającymi im prawdopodobieństwami teoretycznymi znajdującymi się w tabl. 9 można wyciągnąć wniosek, że w 2011 r. dla osób mieszkających w mieście praca najemna oraz praca na własny rachunek poza rolnictwem były stosunkowo częściej głównym źródłem utrzymania niż dla osób mieszkających na wsi. Z kolei dla osób mieszkających na wsi praca na własny rachunek w rolnictwie była stosunkowo częściej głównym źródłem utrzymania w porównaniu z tą grupą osób, które mieszkały w mieście.

PORÓWNANIE WPŁYWU MIEJSCA ZAMIESZKANIA NA ROZKŁAD ŹRÓDEŁ UTRZYMANIA LUDNOŚCI POLSKI W LATACH 2002 I 2011

We wcześniejszych rozdziałach opracowania udowodniono, że na zróżnicowanie Polaków pod względem występowania poszczególnych źródeł utrzymania w 2002 r. i w 2011 r. statystycznie istotny wpływ miało miejsce zamieszkania. Warto więc odpowiedzieć na pytanie, jak silna była ta zależność oraz czy zależność istniejąca w 2011 r. uległa w stosunku do 2002 r. zwiększeniu czy może zmniejszeniu. W tym celu należy określić siłę występujących zależności dla lat 2002 i 2011, a następnie dokonać stosownych porównań.

Do określenia siły zależności posłużył współczynnik V Cramera. Wartość tego współczynnika wyniosła w latach: 2002 — 0,32; 2011 — 0,25.

Można zatem stwierdzić, że miejsce zamieszkania w znacznym stopniu rzutowało na to, jakie było główne źródło utrzymania mieszkańca Polski tak w 2002 r., jak i w 2011 r. Należy jednak zauważyć, że fakt zamieszkiwania

przez daną osobę w mieście czy na wsi w mniejszym zakresie wpływał na zróżnicowanie mieszkańców Polski pod względem głównych źródeł utrzymania w 2011 r. niż było to w roku 2002. Dowodem na to jest niższa wartość współczynnika V Cramera obliczona dla 2011 r. od jego wartości obliczonej dla roku 2002.

W tabl. 12 pokazano różnice między prawdopodobieństwami empirycznymi i odpowiadającymi im prawdopodobieństwami teoretycznymi w 2002 r., a w tabl. 13 różnice między prawdopodobieństwami empirycznymi i teoretycznymi w 2011 r. Z kolei w tabl. 14 porównano poszczególne składniki statystyki testowej χ^2 otrzymane dla 2002 r. z analogicznymi wielkościami obliczonymi dla 2011 r.

TABL. 12. RÓŻNICE MIĘDZY PRAWDOPODOBIEŃSTWAMI EMPIRYCZNYMI I TEORETYCZNYMI W 2002 R. W P.PROC.

Wyszczególnienie	Główne dochody					Suma
	z pracy			ze źródeł niezarobkowych	z własności	
	najemnej	na własny rachunek poza rolnictwem	na własny rachunek w rolnictwie			
Miasta	3,317	0,701	-3,641	-0,389	0,013	0,000
Wieś	-3,317	-0,701	3,641	0,389	-0,013	0,000
Suma	0,000	0,000	0,000	0,000	0,000	0,000

Źródło: opracowanie własne na podstawie tabl. 3 i 4.

TABL. 13. RÓŻNICE MIĘDZY PRAWDOPODOBIEŃSTWAMI EMPIRYCZNYMI I TEORETYCZNYMI W 2011 R. W P.PROC.

Wyszczególnienie	Główne dochody					Suma
	z pracy			ze źródeł niezarobkowych	z własności	
	najemnej	na własny rachunek poza rolnictwem	na własny rachunek w rolnictwie			
Miasta	1,995	0,439	-2,567	0,145	-0,012	0,000
Wieś	-1,995	-0,439	2,567	-0,145	0,012	0,000
Suma	0,000	0,000	0,000	0,000	0,000	0,000

Źródło: opracowanie własne na podstawie tabl. 8 i 9.

TABL. 14. PORÓWNANIE SKŁADNIKÓW STATYSTYKI TESTOWEJ CHI-KWADRAT W LATACH 2002 I 2011

L a t a	Główne dochody					Suma
	z pracy			ze źródeł niezarobkowych	z własności	
	najemnej	na własny rachunek poza rolnictwem	na własny rachunek w rolnictwie			
Miasta						
2002	97379	33114	765261	1205	501	897461
2011	31723	12211	540126	205	354	584620

**TABL. 14. PORÓWNANIE SKŁADNIKÓW STATYSTYKI TESTOWEJ CHI-KWADRAT
W LATACH 2002 I 2011 (dok.)**

L a t a	Główne dochody					Suma
	z pracy			ze źródeł niezarobko- wych	z własności	
	najemnej	na własny rachunek poza rolnictwem	na własny rachunek w rolnictwie			
Wieś						
2002	162578	55285	1277626	2012	836	1498337
2011	52699	20285	897258	340	588	971171

Źródło: opracowanie własne na podstawie tabl. 6 i 11.

Porównanie wyników znajdujących się w tabl. 12 i 13 pozwala na wyciągnięcie wniosku, że w przypadku każdego z prawdopodobieństw empirycznych miało ono wartość bliższą prawdopodobieństwu teoretycznemu w 2011 r. niż było to w roku 2002. Z kolei na podstawie tabl. 14 można stwierdzić, że wszystkie składniki statystyki χ^2 obliczonej dla 2011 r. były dużo niższe od odpowiadających im składników statystyki χ^2 odnoszącej się do roku 2002. Nie ulega więc wątpliwości, że na przestrzeni dziewięciu lat różnice między mieszkańcami miast i wsi w Polsce w zakresie tego, czy dochody ich pochodzą z pracy najemnej, z pracy na własny rachunek poza rolnictwem, z pracy na własny rachunek w rolnictwie, ze źródeł niezarobkowych czy z własności uległy znacznemu zatarciu.

Podsumowanie

Głównym celem artykułu jest udzielenie odpowiedzi na pytanie, jak zmienił się wpływ miejsca zamieszkania na rodzaj głównego źródła utrzymania mieszkańców Polski między latami 2002 i 2011. Przeprowadzone badanie pozwoliło sformułować następujące wnioski:

- w 2002 r. i w 2011 r. istniała w Polsce statystycznie istotna zależność między tym, czy osoba mieszka w mieście czy na wsi a tym, jakie jest jej główne źródło utrzymania. Pobieranie dochodów z pracy najemnej bądź z pracy na własny rachunek poza rolnictwem było typowe dla mieszkańców miast, natomiast pobieranie dochodów z pracy na własny rachunek w rolnictwie było charakterystyczne dla mieszkańców wsi;
- wpływ miejsca zamieszkania na rozkład głównych źródeł utrzymania ludności Polski w roku 2011 był zdecydowanie mniejszy niż w roku 2002;
- różnica między udziałem ludności miast, dla której praca najemna była głównym źródłem utrzymania a udziałem ludności wsi, dla której praca najemna była głównym źródłem utrzymania uległa na przestrzeni dziewięciu lat znacznej redukcji. Prawidłowość taką zaobserwowano nie tylko w przy-

padku pracy najemnej, ale również pracy na własny rachunek poza rolnictwem i na własny rachunek w rolnictwie, własności oraz źródeł niezarobkowych.

W toku przeprowadzonych badań udowodniono, że w 2011 r. struktury ludności w mieście i na wsi — pod względem głównych źródeł dochodów — mniej się od siebie różniły niż można było to zaobserwować w 2002 r. Hipotezę badawczą mówiącą, że fakt mieszkania w mieście albo na wsi w coraz mniejszym zakresie determinuje rozkład mieszkańców Polski pod względem rodzajów posiadanych źródeł utrzymania zweryfikowano pozytywnie. Pozwoliło to na stwierdzenie, że polska wieś i polskie miasta na przestrzeni rozpatrywanych dziewięciu lat uległy pod względem rozpatrywanej cechy znacznemu ujednoliceniu.

dr Anna Turczak — Zachodniopomorska Szkoła Biznesu w Szczecinie, **dr Patrycja Zwiech** — Uniwersytet Szczeciński

LITERATURA

- Aczel A. D. (2000), *Statystyka w zarządzaniu*, PWN, Warszawa
- Buga J., Kassyk-Rokicka H. (2008), *Podstawy statystyki opisowej*, Wyższa Szkoła Finansów i Zarządzania w Warszawie
- Józwiak J., Podgórski J. (1995), *Statystyka od podstaw*, Polskie Wydawnictwo Ekonomiczne, Warszawa
- Kot S., Jakubowski J., Sokołowski A. (2007), *Statystyka. Podręcznik dla studiów ekonomicznych*, Centrum Doradztwa i Informacji „Difin”, Warszawa
- Krysicki W., Bartos J., Dyczka W., Królikowska K., Wasilewski M. (2003), *Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach. Część II*, PWN, Warszawa
- Kukuła K. (2003), *Elementy statystyki w zadaniach*, PWN, Warszawa
- Ludność i gospodarstwa domowe. Stan i struktura społeczno-ekonomiczna 2002* (2003), GUS
- Ludność i gospodarstwa domowe. Stan i struktura społeczno-ekonomiczna 2011* (2013), GUS
- Ostasiewicz S., Rusnak Z., Siedlecka U. (1995), *Statystyka. Elementy teorii i zadania*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu
- Podgórski J. (2005), *Statystyka dla studiów licencjackich*, PWE, Warszawa
- Pułaska-Turyńska B. (2005), *Statystyka dla ekonomistów*, Centrum Doradztwa i Informacji „Difin”, Warszawa
- Zeliaś A. (2000), *Metody statystyczne*, PWE, Warszawa

SUMMARY

The aim of the study is to examine the strength of the association between the place of residence of the Polish population and their main source of income on the basis of the last two censuses carried out in 2002 and 2011. The study used Pearson chi-square test of independence and Cramer's V coefficient to assess

the relationship between the changes in residence place and the main source of income, which occurred between the two last censuses period. The analysis showed that the difference between urban and rural areas, depending on the source of income has been over nine years significantly reduced. This regularity was observed by comparisons between the study years the shares of the population, which wage labor was the main source of income, as well as the shares of the population, the main source of income was a self-employed outside agriculture, self-employment in agriculture, property or non-commercial sources.

РЕЗЮМЕ

Целью статьи является обследование прочности связи между местом проживания населения Польши и их главным источником средств на содержание на основе двух последних всеобщих переписей населения проводимых в 2002 г. и в 2011 г. В обследовании был использован критерий независимости хи-квадрат Пирсона и коэффициент V Крамэра для оценки изменения зависимости между местом проживания и главным источником доходов, которая имела место в межпереписный период.

Проведенный анализ показал, что разница между городом и деревней в зависимости от источников дохода за 9 лет значительно уменьшилась. Эта точность наблюдается при сопоставлении в обследуемых годах доли населения, которого наемный труд был главным источником средств на содержание, а также доли населения, которого главным источником средств на содержание является самозанятость кроме сельского хозяйства, самозанятость в сельском хозяйстве, частная собственность или работа без зарплаты.